accomplish the conversion. Another NPRM might need to be promulgated or, at a minimum, an "intent to convert" notice – with appropriate opportunity for comment – would have to occur. Given the above, and the tenuous need for the "guidance" the Commission proposes in a Section 251/252 context, the Commission and affected industry participants would be better off with the "guidance" being deferred to a later date, perhaps in a different form.[49]

Like Commissioner Furchtgott-Roth, U S WEST ultimately must cast its vote against the current formulation of the Commission's "proposed rules," believing that the framework that has lead to their promulgation is confused and confusing, and calculated to lead to additional confusion for state commissions and affected companies in the future. The procedural "strangeness" of the proposal, as well as its possible impacts on contract negotiations and state regulatory resolutions of those issues is certain, as Commission Furchtgott-Roth predicts, to lead to further litigation if promulgated.

---

[49] Compare Dissenting Statement of Furchtgott-Roth at 4 (referencing a work product more streamlined ("no more than a few" items) and resembling something along the lines of a "Commission white paper"), 6-7. And see U S WEST's references herein to a statement of "principles."

Moreover, it is not clear to U S WEST what the Commission might add, through a "white paper," to the "parity" obligations associated with Sections 251 and 252 already announced in its First Report and Order. To the extent the Commission seeks to provide "guidance" regarding Sections 271/272 requirements, a "white paper" might prove helpful, provided it did not seek to insinuate obligations outside of those Sections 251/252 matters left to the Commission's jurisdiction into the regulatory mix.

C. The Objectives Of "Non-Legally Binding Guidance" Is Oblique.

The "goals" associated with the Commission's proposals remain oblique. For example, the Commission proposes that its non-legally binding "guidance" would provide ILECs with an ability "to use the performance measurements as evidence of compliance with their relevant statutory obligations in order to counter allegations of noncompliance."[50] However, the Commission does not identify what statutory or regulatory compliance it is addressing.

Frankly, ILECs can achieve compliance with Sections 251 and 252 right now by pointing to compliance with contractual agreements and by pursuing the arbitrations and mediations associated with state authority.[51] And – while U S WEST cannot speak for all carriers – we believe that carriers are more likely to negotiate additional reporting measures than accede to their being mandated by federal regulatory authorities. Thus, as addressed above, the Commission's current proposed regulatory model might well prove merely to provide fodder for litigation.[52]

V. MATTERS PERTAINING TO THE COMMISSION'S SPECIFIC PROPOSALS.

Because U S WEST will undoubtedly have to address any proposed

---

[50] NPRM para. 5.

[51] Dissenting Statement of Furchtgott-Roth, at 2 ("Even the most casual of conversations with any State Commissioner reveals that OSS issues are closely monitored and addressed by the States.").

[52] See id. at 5.

measurements adopting by this Commission as guidance to the States in both contract negotiations and state arbitrations, mediations and perhaps broader state-initiated rulemaking proceedings, we do not address the Commission's specific proposed measurements here. Suffice it to say that the Commission's proposals contain proposed measurements that U S WEST does not currently measure and which it has not bound itself by contract to measure.

In some circumstances, U S WEST's failure to measure performance with respect to a specific item has to do with our current inability to do so (i.e., no measurement capability exists, either because the item was not previously measured or because measurement of the performance involves a service U S WEST does not render). In other cases, we believe that the costs of creating the measurement capability outweigh the benefits or that a somewhat differently formulated measurement is a better measurement than that proposed by the Commission (because its output information is itself not as subject to interpretation as would be the case utilizing the Commission's proposed measurement).

While we do not address the specific proposed performance measurements, below we do address those matters of a "general" nature that the Commission seeks comment on with respect to its proposal.[53] We also address the matter of the scope of reporting requirements and obligations as well as the evaluation of such reports and the information contained therein.

---

[53] NPRM para. 27.

A. Matters of General Applicability.

1. Balance Between Burdens and Benefits.

As a general matter, it is fair to say that to the extent a new measure or report can be obtained from existing internal procedures, the proposed rules would not place an undue burden upon U S WEST from a practical perspective. On the other hand, to the extent the Commission's proposals require the collection of data or measurement of things that U S WEST does not currently measure, there will clearly be a cost burden associated with the creation of the measure. The creation of an obligation to measure or report anything beyond that which U S WEST has either agreed to provide or been required to provide under lawful state authority would create inappropriate "burden."

Additionally, as a general matter, U S WEST believes that fewer, rather than more, performance measurements (with the Commission's proposal representing an elaborate "more" approach) is sufficient to provide adequate information upon which a non-discrimination verification can be accomplished. The greater the number of measures, the greater the burden on ILECs to collect and report, often with increasingly marginal benefit.

Finally, the Commission poses the possibility that certain reporting methodologies might be structured in such a manner so as to create a "safe harbor"

for ILECs with respect to challenges associated with nondiscriminatory activity.[54] While U S WEST general supports the intellectual notion of safe harbors, and engages in vigorous advocacy at the state level regarding the "meaning" of statistical variations and their "significance,"[55] it is unclear what the Commission believes it can add to the "safe harbor" notion.

Particularly in light of the fact that the Commission curiously does not craft its proposals around 271 issues and the role that a "safe harbor" notion might play in such proceeding, U S WEST believes the "benefit" associated with trying to structure a legally binding federal "safe harbor" is far outweighed by the burden associated with the Commission's unduly detailed proposals. For example, how "safe" would a federal safe harbor be, given the deference the Telecommunications Act provides to state authorities as moderators of private contractual negotiations and as first-line decision makers with respect to Section 271 filings? Since under the Commission's proposal (and the NARUC's Resolution), states are free to deviate from any federally-proposed measurements, the notion of any federal safe harbor would undoubtedly prove illusory. And, that illusion would surely breed only further confusion, contention and – undoubtedly -- litigation.

For all of the above-stated reasons, the Commission's overall proposal is one

---

[54] NPRM para. 21 (addressing a possible use of statistical analysis, wherein if an ILECs performance fell within a predetermined range, an ILEC would be free from challenges of discrimination).

[55] U S WEST's position is that a "significant" difference in performance requires both a statistical and operational significance. In the attached Appendix A, incorporated herein by this reference, U S WEST responds to the statistical sampling issues the Commission raises in Appendix B of the NPRM.

far more suggestive of burden than benefit.

2. Geographic Level For Reporting.

The Commission seeks comment upon the appropriate geographic level for reporting.[56] U S WEST supports a statewide geographic reporting area. U S WEST's service region is unique among the ILECs in that it is vast and in many areas sparsely populated. Particularly in light of the nature of U S WEST's geographic region, we believe a statewide reporting level is the most appropriate with respect to minimizing expense.

3. Scope Of Reporting.

The Commission seeks comment on the appropriate scope of reporting, proposing that performance measurement reports be provided on the ILEC's performance as to its own retail customers, any of its affiliates that provide local exchange service, competing carriers in the aggregate, and individual competing carriers.[57] As a general matter, U S WEST would not have a problem with a performance measurement and reporting structure reflecting the reporting scope proposed by the Commission.[58]

---

[56] NPRM, para. 38.

[57] Id. para 39.

[58] Throughout the discussion in this Section, it should be understood that when U S WEST says it would "not have a problem" with a Commission proposal or finds it non-burdensome (or expresses similar ideas in different words), we do not mean to suggest that we believe the Commission has the authority to impose the measurement. What we mean to suggest is that such might be a proper item of negotiation and, if agreed to or ordered through a state-authorized arbitration or mediation, U S WEST might not pursue a judicial appeal.

However, we believe that any intelligent discussion on the "scope" of reporting must begin with the standard which the reporting is directed toward. And, we take issue with the Commission's proposal to include ILEC affiliates providing local exchange services as one of the relevant entities with respect to the reporting of performance measurement data, where the reporting involves performance measurements for other than interconnection.

a. The Standard Associated With The Reporting.

As the Eighth Circuit made clear, ILECs are required to provide network interconnection, UNEs and resale utilizing the network they have.[59] They are not required to build superior networks or systems to serve the needs of their competitors. A necessary connotation from the Court's analysis, is that discrimination does not occur when an ILEC stands firm in its position that it is required to collect information and report that information only with regard to functions that it generally has collected data and measured for itself or to support an absence of discrimination as between and among the CLECs it serves.[60]

---

[59] Iowa Utilities Bd. v. F.C.C., 120 F.3d at 812.

[60] As the Commission's proposed measurements make clear, certain of the measurements are designed to capture information only with respect to CLECs, since there is no "LEC analog". NPRM n.154. See, e.g., Appendix A, A2 and A4-A10 (where certain comparisons are between CLECs and ILECs and for others the measurement and reporting is confined to CLECs), A3, A14, A17. Such data collection and reporting, then, serves to prove the existence/non-existence of discrimination within the CLEC class, not between CLECs and the ILEC.

b. Reporting With Respect To Affiliate Transactions.

The Telecommunications Act references nondiscrimination obligations with respect to LEC affiliates only in Section 251(2)(c)(2)(C), which relates to interconnection. There is no reference made to ILEC performance vis-a-vis its affiliates with respect to accessing UNEs, resale, or collocation.[61] Therefore, while the inclusion of a LEC affiliate that provides "local exchange service" as one of the "comparison" entities might be appropriate for interconnection data collection and reporting, it is not appropriate with respect to performance measurements associated with UNEs (including OSSs, OS and DA). Unless a LEC agrees to include such an affiliate in its performance measurement/reporting structure, regulatory authority should not mandate the inclusion.[62]

4. Relevant Electronic Interfaces.

The Commission seeks comment on its tentative conclusion that reporting only with respect to transactions accomplished through electronic interfaces should be required,[63] since this is the primary method by which ILECs access their own

[61] 47 U.S.C. § 251.

[62] While there might be some need to measure performance of a BOC's Section 272 affiliate, to allow for a nondiscrimination finding with respect to the items outlined in Section 272(c)(1) – which includes the provision of "services" – this would be the exception to the absence of an affiliate reporting requirement, not the norm.

[63] This aspect of the Commission's discussion seems focused on OSSs, and U S WEST will respond in that manner.

internal systems and databases.[64] U S WEST supports the Commission's conclusion, since confining any reporting obligation to electronic interfaces allows carriers to capitalize on the kind of "electronic coding" the Commission hopes to tap into.[65]

With respect to the types of electronic interfaces regarding which reporting might be required, the Commission specifically notes that "LECs provide several types of electronic interfaces, such as a GUI-based interface and an EDI-based interface."[66] Because of the potential multiplicity of electronic interfaces, the Commission inquires whether ILECs should report performance as to all formats of electronic interfaces or some subset thereof.[67]

As a general matter, U S WEST believes that the contract negotiation process will seek to accord CLEC parties to interconnection contracts with performance information associated with the type of interface they utilize – be it GUI or EDI. While all relevant information cannot now be captured or reported (at least in U S WEST's case), the goal is to define the appropriate performance measures for the specific contracting parties' needs.

---

[64] NPRM para. 40.

[65] Id. para. 42.

[66] Id. para. 41 (footnotes omitted).

[67] Id. paras. 41-42.

B. Reporting Procedures.

With respect to reporting procedures, the Commission seeks comment on three items: (1) the recipients of reports; (2) the frequency of the reports; and (3) auditing procedures.[68] U S WEST addresses each of these items below.

1. Receipt Of Reports.

The Commission notes that "the main purpose of . . . performance reports is to permit competing carriers to determine whether they are obtaining access consistent with the requirements of section 251."[69] For that reason, the Commission tentatively concludes that "only those carriers that already obtain services or facilities from the [ILEC] through an interconnection agreement, or under a statement of generally available terms, should have the opportunity to receive reports.[70]

U S WEST agrees with the tentative conclusion imposing a "doing business" requirement on CLECs prior to their being able to demand reports.[71] For that reason, we agree with the Commission's tentative conclusion that reports should only be provided to CLECs that have activity in the reporting period for a particular service to which they subscribe. There should be no requirement that an ILEC

---

[68] Id. para. 104.

[69] Id. para. 106.

[70] Id.

[71] In this regard, U S WEST would add the requirement that the "doing business" relationship be free of any breaches of the receiving carrier's obligation to the ILEC. Thus, any CLEC otherwise eligible to receive a report should not receive such a report if the CLEC is not current in its obligations, including payment for charges, joint planning and forecasting requirements, etc.

report to a CLEC data related to access to UNEs, for example, if that particular CLEC has not ordered the particular UNEs during the reporting period.

However, we do not believe that federal regulatory intervention is necessary to either define the measurements in such reports or to entitle CLECs to receive them. CLECs in U S WEST's territory already know how to negotiate for such reports (if they are interested in receiving them – and some carriers are not interested).[72] And, they know how to use them (i.e., "whether to try to resolve the problem through discussions . . . or whether some other action, such as filing a complaint, is required.").[73]

U S WEST also agrees with the Commission that State commissions are best able to determine their own need for reports. It has been U S WEST's experience that states, generally, prefer to have reports go to directly-affected CLECs, requesting such reports only if independently deemed necessary.

The Commission also seeks comment on the establishment of a "clearinghouse" as a central repository for ILEC reports.[74] Because the entitlement to have ILECs engage in performance measurements, and to the ultimate reports outlining the results of those measurements, is fundamentally one of contract, U S WEST opposes the idea that information associated with such reporting be used to populate some kind of "clearinghouse."

---

[72] This fact supports the Commission's tentative conclusion that ILECs should only provide reports to carriers requesting them. NPRM para. 106.

[73] Id.

[74] Id. para. 109.

Just as with other matters addressed in the interconnection contracts, the use of the performance measurement information provided to the CLECs is negotiated. U S WEST negotiated contracts with CLECs contain provisions which address information provided to CLECs through reporting vehicles. Those provisions restrict the use of the information to matters addressed in the specific contract with the specific CLEC. They also require that access to reported data be limited to internal use by those directly involved in the non-discrimination verification process for that CLEC. Specifically, performance measurement data contained within the reports is prohibited from being shared internally with marketing,[75] public relations and other departments within the CLEC not directly involved in the underlying interconnection or service relationship.

The establishment of a publicly-available clearinghouse of information would be at odds both with the privity of contract allowed to negotiating carriers and the need to manage contention at the level closest to the negotiating process. As indicated by Dr. Carnall in the attached Appendix, attempting to use statistical processes for cross-state or cross-region comparisons adds another level of complexity to an already complex endeavor.[76]

Furthermore, given that States are required to base their determinations of Section 271 compliance on information associated with services provided in their

---

[75] For example, if there were a small number of CLECs, it might be possible for one CLEC to "break out" its information, and attempt to use the disaggregated information to "market" its service performance. U S WEST has prohibited such use.

[76] Appendix A at 8.

states, there is no need for them to be "sharing" information (other than that which they might ultimately share informally).

Finally, any clearinghouse would undoubtedly require some source of funding. ILECs have already been saddled with tremendous and extraordinary expenses associated with compliance with Sections 251 and 252 of the Telecommunications Act. They should not be burdened further by costs associated with the establishment of a "clearinghouse."

2. Frequency Of Reports.

U S WEST supports providing reports on a quarterly basis and opposes the provision of reports on a more frequent basis (such as monthly).[77] Not only would a reporting interval more frequent than quarterly be expensive and burdensome on ILECs, but it would also provide less meaningful data. Quarterly reporting allows for a greater amount of information to be observed over a longer period of time, lessening the impact of transient results that might occur – results that could inappropriately skew the overall accuracy of the data reported.[78]

Moreover, U S WEST suggests that any rules governing reporting should have built-in sunset provisions, preferably two years from the date of enactment. Given the deregulatory thrust of the 1996 Act, and the expectation that some RBOCs will have secured Section 271 relief (to the extent such is legally required) within the next two years, a sunset provision is appropriate.

---

[77] NPRM para. 112 (where the Commission observes that monthly reporting could involve "significant costs," a conclusion with which U S WEST agrees).

[78] See Appendix A at 3.

3. Auditing Requirements.

Like other areas dealing with performance measurements, the matter of "audits" of raw data and processes are routinely a matter of contract negotiation. Generally, audit rights that are negotiated are not confined solely to OSS performance, for example, but are an integral part of the overall contract relationship. For this reason, U S WEST opposes inclusion of any auditing mechanism in the Commission's proposed rules. For those CLECs with existing contracts who desire performance measurement reports, audit rights already exist.

C. Evaluation Of Performance Measurements.

The Act requires non-discriminatory provisioning and access to service and elements. The appropriate use of the performance measures is to assist in verifying the absence of unlawful discrimination. To the extent that there are valid comparisons between a CLEC result and an ILEC result, factual reporting of "results" are just that: facts.

Those facts may be indicative of possible discrimination or may be the result of some other cause. If, after applying a statistical significance test,[79] results between U S WEST and a CLEC are substantially the same, the result would suggest that no discrimination exists. If such results are not substantially the same, further investigation might be warranted, if desired by the CLEC.

[79] NPRM para. 121 (where the Commission seeks comment upon the appropriate use of any statistical testing process).

Whereas a "positive" showing of "sameness" as between an ILEC and a CLEC warrants some deference (in the form of a "safe harbor," for example),[80] an alleged "negative showing," based on statistical variations or deviance should not be entitled to a presumption that discrimination has occurred. To further explain the issues around statistical analysis, U S WEST provides an Appendix in which Dr. Michael Carnall explains (1) the necessity for implementation of a statistical test; and (2) the proper use of the result. Simply, if a statistical test, such as the two-tailed Z test advocated by U S WEST, designed to a 99% confidence level, shows a material difference in results between U S WEST and a CLEC, then that result is statistically valid and may be indicative of possible discrimination.

In essence, the application of the Z test serves to help define a"threshold" for determining if further inquiry is necessary or desired. For results that fall within the range defined by the 99% confidence level, the test would be assumed to establish non-discrimination and no further action would be required (i.e., a form of "safe harbor" approach).

VI. OTHER ISSUES RAISED BY PETITIONERS.

The Commission seeks comment upon several matters raised by Petitioners which the Commission has tentatively determined not to address in this proceeding.[81] Below, U S WEST provides brief comment on each of the areas regarding which the Commission seeks comment.

---

[80] See pages 24-25, supra.

[81] NPRM paras. 124-30.

A. Performance Standards.

Performance standards should not be adopted. U S WEST agrees with the Commission that promulgation of performance standards requires some grounding in "historical experience to ensure that such standards and fair and reasonable."[82] Lacking such experience, we agree that it is premature to proceed toward any type of standard development under the mantle of federal regulatory authority.

B. Technical Standards.

U S WEST agrees with the Commission's approach to the development of technical standards for OSS, i.e., to leave the matter in the hands of industry committees that are currently engaged in working through the issues.[83] Through industry fora, much progress has been and continues to be made in developing appropriate OSS interfaces.

As the Commission notes, such industry forums are open and have the resident technical expertise to grapple with technical issues that are not easily resolved *via* the formal federal rulemaking process. Furthermore, standards development involves not just the establishment of the standard but the ultimate decision to deploy the standard – a downstream business decision.

Because of the dual-pronged nature of standards work, U S WEST opposes the idea of Commission-mandated pre-determined "implementation dates," based on

---

[82] Id. para. 125.

[83] Id. paras. 126-29.

the final development of a standard through an industry forum.[84] This type of regulatory intervention is totally unnecessary, overreaching and inappropriate. It ignores the material issues of commercial analysis after the development of the standard itself. The Commission should not preemptively supplant the commercial analysis aspect of the standards process and should reject the notion of "automatic implementation" of adopted standards.

C. Enforcement Mechanisms.

U S WEST also agrees with the Commission's determination that it is premature to propose enforcement mechanisms.[85] Clearly, given the nature of the Commission's approach to the proceeding, i.e., the establishment of non-legally binding "guidelines," "enforcement issues" are best left to the states.

VII. CONCLUSION.

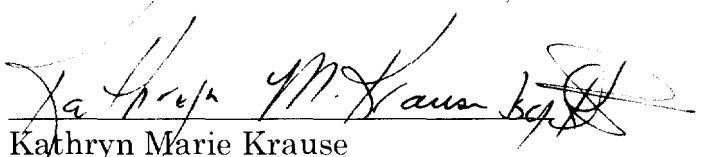
For all of the above reasons, the Commission should terminate the instant rulemaking proceeding. The matters addressed in the NPRM are best left to contract negotiations and state regulatory oversight. To the extent the Commission believes "guidance" to the states is necessary regarding OSS performance and Section 271 relief, it should craft its "guidance" in the form of generally-applicable "principles" (similar to the principles associated with number portability). Beyond

[84] Id. para. 129.

[85] Id. para. 130.

that, no federal regulatory authority should be extended over ILEC performance measurements or reporting.

U S WEST COMMUNICATIONS, INC.

By: 

Kathryn Marie Krause
Suite 700
1020 19th Street, N.W.
Washington, DC 20036
(303) 672-2859

Its Attorney

Of Counsel,
Dan L. Poole

June 2, 1998

Appendix A
Comments of U S WEST Communications, Inc.
Comments of Michael Carnall on Statistical Issues of Detecting Differences in Service Quality

I. INTRODUCTION

A. Qualifications

I, Dr. Michael Carnall, am currently a Senior Economist at LECG Inc. At LECG, my work in telecommunications has focused on the analysis of economic and cost modeling as well as service quality issues. I have also done work in the electric power industry, where I was involved in projects concerned with power quality and the optimization of transmission grid security in a deregulated industry.

Prior to joining LECG in 1996, my professional experience included eight years of involvement in the measurement and analysis of product quality and reliability at Caterpillar Inc. in Peoria, Illinois. In connection with that position, in 1982 I developed a system for tracking and evaluating the quality and reliability of prototype products. I later developed a method of directly applying field reliability analysis to the redesign of product components and coordinated the use of the method in the design of the powertrain of Caterpillar's largest track type tractor. As Senior Reliability Analyst, I conducted seminars on the theory, interpretation and use of field quality and reliability measurements throughout the U.S. and at subsidiary plants in Leicester, England; Gosselies, Belgium; Geneva, Switzerland; Grenoble, France; Tokyo, Japan; Sao Paulo, Brazil; and Melbourne, Australia. As part of my responsibility to train new reliability analysts, I taught short courses in reliability analysis as required. During the 1990 corporate restructuring, I participated in the development of product quality and reliability measures used in division incentive schemes. I was also responsible for developing and administering the allocation of warranty expenses to each of the newly formed divisions.

When I left Caterpillar in 1991 to enter the doctoral program at the University of Illinois, I was retained by the Company as a consultant to develop a system which gathered production quality data, matched that data with the subsequent field information and examined the correlation between the two experiences. As a teaching assistant at the University of Illinois, I taught undergraduate statistics in the College of Commerce.

My academic credentials include a Doctor of Philosophy degree in Economics from the University of Illinois at Urbana-Champaign (1996). Prior to that, I earned Bachelors (1977) and Masters (1986) degrees in Civil Engineering from Bradley University.

B. Executive Summary

Any test designed to discover the presence of discrimination in the delivery of service operations must use statistical methods to properly account for the random nature of these operations. Although Incumbent Local Exchange Carriers (ILECs) may have been monitoring the quality these operations for some time, the data has not been used in the ways proposed in the Commission's NPRM. Nor have the statistical tests proposed been utilized with these data in exactly the ways proposed. In its NPRM the Commission has requested comments on the applicability and limitations of statistical tests for detecting the presence of discrimination in telecommunications service operation.[1] In the following, I address the theoretical bases for these tests and contrast them to the reality of the data upon which they will be brought to bear. Then, using Monte Carlo methods, I assess the effect of probable deviations of the actual data from the ideal conditions assumed in the development of the tests. I find that the most common parametric tests for detecting whether two samples are from the same population, the *t* and *Z* test, are sensitive, in varying degrees, to both sample size and to the distribution of the underlying data. These sensitivities affect the realized probability of a false finding of discrimination as well the ability to detect various types of differences in the underlying populations.

From these analyses I conclude that while these statistical tests can provide invaluable guidance in the detection of discrimination they must be applied very carefully. Test parameters such as the confidence level must be set with proper regard to the characteristics, including sample sizes and distribution, of the data being tested. Failure to do so will lead to erratic results and either an excessive number of false alarms or tests which are insensitive to the presence of discriminatory service quality.

C. Service intervals are random

Efforts to assure substantially similar service must address operations which depend upon human intervention as well as automated mechanical and electronic operations. The "quality" of these operations, as measured, for example, by the time required to perform them, is influenced by

[1] FCC-NPRM 98-72. Appendix B.

unforeseeable external events and conditions. As a result their measured "quality" is characterized by a range of measurements and associated probabilities, what statisticians generally refer to as a distribution. For example, 75 percent of all of all residential installation changes may be accomplished within one day, 80 percent within two days and 90 percent within five days. Such a distribution, taken over a long period, describes the "population" from which all equivalent operations are taken. It accurately reflects the time required to accomplish the operation and also the influence of all random external factors that might affect the operation.

It is important to note that at least some of the external factors influencing quality are not completely random, in the true sense of the word. Such things as weather, location of customer, number of other orders pending, day of week, day of month and season of the year may have an influence on the time required to complete an operation. If a large percentage of all orders are placed near the end of the month, the time required to complete an order submitted during that period might well higher than that for a similar order submitted at the beginning of the month. Over a long period, for a single provider, the influence of such factors is conveniently included as random variation. When monitoring intervals over as short a period as a month, controlling for their influence may be necessary in order to understand what may appear to be wide variation in service quality measurements.

Given the influence of the external factors described above and the varied nature of Competitive Local Exchange Carriers (CLECs), it is very likely that these external factors will influence the operations ordered by a particular CLEC differently from those ordered by other CLECs, as well as from those provided to the ILEC itself. If customer location (rural versus urban, for example) significantly increased the time required to establish or modify service, it would not be unexpected to observe higher installation intervals for a CLEC that targeted only rural customers. In this case it would no longer be valid to assume that these service operations were taken from the overall distribution of installations for the ILEC. One would more properly assume that they were drawn from the "marginal" distribution of the ILEC which contained only rural customers. The marginal distribution would properly describe the ranges and probabilities of installation intervals, given that the customer that was located in a rural area.

Telecommunications operations can, of course, also be influenced by biases built into mechanical and electronic systems and the incentives of their human operators. Such biases can systematically increase or decrease the quality of service as observed in the service quality measurements. It is these inadvertently or intentionally introduced biases that the measurements of service quality should be designed to detect.

1. Impossible to detect, with certainty, differences in quality

Even after substantial work has been done to adjust for external factors and to determine the proper marginal distribution to which a CLECs orders should be compared, there will still be a significant amount of random variation remaining in any set of service quality measurements. The presence of this variation makes it **impossible** to determine, with certainty, whether the service operations performed for a CLEC by an ILEC are systematically of a lesser quality than those the ILEC performs for itself or for other CLECs. The best that one can hope to accomplish under these circumstances is to estimate the probability that the quality provided to a particular CLEC is different from that provided by the ILEC in serving its own customers.

It is an important first step to recognize and acknowledge the impossibility of eliminating all error from the determination of equal quality. It is also important to recognize that statistical techniques can provide guidance in the task of systematically estimating the probability that a service operation is being provided with substantially similar quality. These techniques provide, for example, ways to calculate from observed data, estimates of the probability that two sets of service operations were provided from the same population of operations.

2. False findings of discrimination are inevitable

Since it is impossible to determine with certainty whether a service operation is being provided at an equal level of quality, it follows that there will always be some possibility of a false finding of discrimination in the provision of the service. Using statistical techniques one can estimate that there is only a one percent probability that a specific set of observed data could result from non-discriminatory provision of service. Such evidence may at first seem unequivocally damning. Put another way, however, if an ILEC is providing absolutely substantially similar quality service, under such a rule it will be declared guilty of discrimination for one measurement out of one hundred.

3. Tests should provide a low probability of false findings of discrimination

Recognition of the inevitability of false findings of discrimination is an important element in the design of tests for the detection of discrimination in the provision of telecommunications services. It is imperative to the credibility of the test that the probability of making a false finding of discrimination be very small. We are all familiar with the tale of the boy who cried wolf in jest so often that he was ignored and devoured when a real wolf appeared. Any attempt to utilize a

test that too often falsely declares discrimination will inevitably result in the test being ignored and any real discrimination going undetected.

A recent Associated Press story describes a multi-million dollar system designed to detect explosives in air passenger baggage that is going unused because it produces too many "false positives." The system indicates the presence of explosives in baggage when there was nothing but food in the suitcase.[2] Any statistical test for discrimination in service should properly recognize the inevitability of false positives and set the level of false positives appropriately small.

D. Confidence in Test Results

1. Statistical "confidence" is a mathematical construct.

The term "confidence," as used in statistics is associated with a particular set of mathematical constructions based on very specific and confining assumptions. A statement that "the sample data indicate, with 99 percent confidence, that the CLEC data is drawn from a distribution with a lower mean than that from which the ILEC data is drawn" is properly interpreted to mean that where sample data has the characteristics of the current samples and where the underlying data conforms to all of the assumptions upon which the test is based, there is only a one percent probability that the ILEC and CLEC data were drawn from the same population.

As defined by Webster's and as usually understood by laymen, confidence describes "an assurance of mind or firm belief in the truth and reality of a fact," a much broader concept. It is this definition of confidence, one which subsumes the issues of data generation and characteristics as well as the treatment of purely random variation, that we would like to have in the results of any testing procedures which are put into place. When doing critical reliability analyses, one often investigates each and every reported failure to ensure that it accurately represents the event that occurred. Doing so does not alter the statistical confidence level of the analysis but it has a profound affect one's confidence in the results of the analysis.

Statistical methods provide more or less straightforward recipes for the construction of tests with various calculated "confidence levels." The *Z* and *t* tests are good examples of such recipes. Unfortunately it is too often forgotten that many assumptions have been made in order to

---

[2] Stoller, Gary. "False Alarms Plague Airport Bomb Finders." USA Today. May 5, 1998.

facilitate the construction of these tests and the use of standard statistical tables to determine critical test values. Subsumed in the development of these tests are assumptions about the independence of the samples and the distribution of the underlying population from which the sample was drawn. Any deviation from these very specific conditions will distort, often in unknown ways, the inferences drawn from the tests.

(1) Independence of samples

Independence of a sample simply means that the variation in the sample data reflects only random variation. That is, it is not systematically influenced by any external non-random conditions. Almost all real sample data violates this assumption to some degree. A more candid description of most data is that "the sample data is not systematically and differentially influenced by any known and measurable external condition." That is, the data may be systematically influenced by external conditions, but those external conditions affect each sample in the same way and therefore their influence can be considered to be "random."

For data on telecommunications service operations the first description would apply to sample data which has been corrected to remove the influence of all seasonal or temporal variations as well as any systematic variation due to type or location of customer. Clearly such data will be difficult to obtain. The second description would apply to data which has been separated into categories so that each category is similarly influenced by all external factors. This definition acknowledges that complete independence is not a realistic requirement but that by separating the data into reasonably homogenous categories the worst effects of dependence can be mollified.

(2) Distribution

All of the most widely proposed parametric statistical tests are based upon assumptions about the distribution of the underlying population from which the data is drawn. The *t* test is based on the assumption that the underlying population has the characteristics of a normal distribution. The theoretical development of the Chi-square test for equality of sample variances also relies on the normalcy of the underlying population. Although, in theory, the *Z* test does not depend on the shape of the underlying distribution, it is strictly applicable only to infinite samples.

Each of these tests is, to some extent, "robust" to the violation of the assumptions upon which it is based. By that is meant that the tests produce reasonably accurate results even if some of the assumptions, either sample size or nature of underlying distributions, are not met.

Unfortunately the nature and extent of the error that will be introduced when one or more of these conditions is violated depends upon the degree to which the assumptions are violated. Any attempt to accommodate for the error is therefore specific to the data source under consideration.

2. Real confidence in the accuracy of a test must include all factors which might influence the accuracy of the test.

Establishing a test in which there is a high level of confidence, in the broader sense that the term is used by laymen, is not limited to the proper treatment of purely random effects. In order to be truly confident in the results of a statistical test we must do more than simply apply the appropriate recipe.

a) Accuracy of the data in reflecting the underlying phenomenon being measured.

The most basic requirement of the data used as the basis of any test is that it should accurately measure the phenomenon under question. Service intervals for installation of new residential service must accurately reflect the time taken to install residential service. Although that statement may seem a trivial tautology, it should not be taken lightly. Unless all service data is properly coded and recorded by technicians, this requirement will not be met. Residential service may be improperly coded as business, the nature of the operation may be misinterpreted or additional operations may be improperly included. Times and dates can be erroneously entered or read and the opportunity for misinterpretation of instructions is infinite. The confidence in the "reality" of a test result based on any body of data must reflect the confidence that the data does in fact measure what it purports to measure.

b) Conformity of the data with the assumptions on which the test is based.

When faced with data from an unfamiliar source, one must usually make some assumptions about the nature of the distribution of the underlying population. If the distribution does not meet all of the criteria for the test to be used, some assessment must be made about the possible affect of the deviation. But unless there is a better suited option available, the test is often applied as if all of the criteria were met. Since most tests are performed one time only on a limited set of data, the test will provide the best information available. Seldom is an effort made to systematically determine the degree to which the data deviates from the ideal and the extent of the associated distortion of the test results.